

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 21:10:37 / Search time 787.52 Seconds

(without alignments)
9201.802 Million cell updates/sec

Title: US-09-988-971-1_COPY_694_942

Perfect score: 249
Sequence: 1 tggcgtatgagggcctgag.....aggccctgctgaccattac 249

Scoring table:

IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Listing first 45 summaries

GeneB1:*

1: gb_da:*

2: gb_hcg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pac:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_da:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pac:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_hcg_hum:*

31: em_hcg_inv:*

32: em_hcg_other:*

33: em_hcg_mus:*

34: em_hcg_pln:*

35: em_hcg_rtd:*

36: em_hcg_mam:*

37: em_hcg_vrt:*

38: em_sy:*

39: em_hgo_hum:*

40: em_hgo_mus:*

41: em_hgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	249	100.0	737	9	AF290986 Homo sapi
2	249	100.0	786	9	AF290985 Homo sapi
3	249	100.0	786	9	AF290986 Homo sapi
4	249	100.0	1183	6	AF290986 Homo sapi
5	249	100.0	1183	6	AF290986 Homo sapi
6	249	100.0	2415	9	AF290986 Homo sapi
7	249	100.0	2415	9	AF290986 Homo sapi
8	213.8	85.2	1348	10	AF290986 Homo sapi
9	212.2	85.2	1321	10	AF290986 Homo sapi
10	150	60.2	14506	9	AF290986 Homo sapi
11	150	60.2	14583	2	AF290986 Homo sapi
12	123.6	49.6	16849	2	AF290986 Homo sapi
13	122.6	49.2	114096	2	AF290986 Homo sapi
14	106.8	42.9	14583	2	AF290986 Homo sapi
15	100.2	40.2	1926	9	AF290986 Homo sapi
16	100.2	40.2	2015	6	AF290986 Homo sapi
17	100.2	40.2	2015	9	AF290986 Homo sapi
18	100.2	40.2	2044	9	AF290986 Homo sapi
19	100.2	40.2	3701	5	AF290986 Homo sapi
20	100.2	40.2	4073	5	AF290986 Homo sapi
21	99.2	39.8	2105	9	AF290986 Homo sapi
22	97	39.0	675	6	AF290986 Homo sapi
23	97	39.0	675	6	AF290986 Homo sapi
24	97	39.0	1491	9	AF290986 Homo sapi
25	97	39.0	1515	9	AF290986 Homo sapi
26	97	39.0	1530	9	AF290986 Homo sapi
27	97	39.0	1589	9	AF290986 Homo sapi
28	97	39.0	2032	9	AF290986 Homo sapi
29	97	39.0	2032	9	AF290986 Homo sapi
30	97	39.0	2129	9	AF290986 Homo sapi
31	97	39.0	2182	9	AF290986 Homo sapi
32	96	38.6	2298	5	AF290986 Homo sapi
33	95.4	38.3	1527	9	AF290986 Homo sapi
34	90.8	36.5	1533	6	AF290986 Homo sapi
35	90.8	36.5	2041	9	AF290986 Homo sapi
36	89	35.7	1911	6	AF290986 Homo sapi
37	89	35.7	1911	10	AF290986 Homo sapi
38	89	35.7	1911	10	AF290986 Homo sapi
39	89	35.7	1940	10	AF290986 Homo sapi
40	89	35.7	2100	10	AF290986 Homo sapi
41	89	35.7	2102	10	AF290986 Homo sapi
42	89	35.7	2172	10	AF290986 Homo sapi
43	89	35.7	2200	10	AF290986 Homo sapi
44	87.6	35.2	2300	9	AF290986 Homo sapi
45	87.6	35.2	2354	6	AF290986 Homo sapi

ALIGNMENTS

RESULT 1

LOCUS AF290986 737 bp mRNA linear PRI 22-APR-2002

DEFINITION Homo sapiens Src-like adaptor protein-2 splice isoform mRNA, complete cds.

ACCESSION AF290986

VERSION AF290986.1 GI:17351922

KEYWORDS

SOURCE

ORGANISM Homo sapiens.

REFERENCE 1 (bases 1 to 737)

AUTHORS Loreto, M.P. and McGlade, C.J.

TITLE Direct Submission

Pred. No. is the number of results predicted by chance to have a

/db_xref="GI:16797892"
/translation="MSGLSPSRKSLPSPLSSSYGCGEYVMEERSKATANAASPP
AGPAPSLRIGEPITIVSEDDMTVISEVSRBNITSVAVAVSHGLVGLSRE
KAEILLPLPGAGPLRESQTRGRSISVRLSPASMDRIIRYRIHCLDNGMLYI
SPRLTFPSLALVDHYSELADICCLLKEPVLQRAGPLPKDIPLPVTVQRTPLNWK
ELDSLLPSEMAATGESLLSEGLRSEISFYSLNDEAVSLDDA"

BASE COUNT 162 a 234 c 231 g 159 t
ORIGIN

Query Match 100.0%; Score 249; DB 9; Length 786;
Best Local Similarity 100.0%; Pred. No. 1.3e-58;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGCTGATGAGGGCCCTGAGCAGGAGAAAGCAGAGAACTGCTGTTTACTCGGGAAC 60
DB 280 TGGCTGATGAGGGCCCTGAGCAGGAGAAAGCAGAGAACTGCTGTTTACTCGGGAAC 339
QY 61 CCGTAGAGGGCCCTTCTCTATCCGGGAGAGCCCAACAGAGAGGCTCTTACTCTGTCA 120
DB 340 CCGTAGAGGGCCCTTCTCTATCCGGGAGAGCCCAACAGAGAGGCTCTTACTCTGTCA 399
QY 121 GTCCGCTCAGCCGCTTCATCTCGGAGCCGATCAGACTACAGATCCACTGCTT 180
DB 400 GTCCGCTCAGCCGCTTCATCTCGGAGCCGATCAGACTACAGATCCACTGCTT 459
QY 181 GACATGAGCTGCTGCTGACTCTCAACCCGCTTCACTTCCCTCACTCCAGGCTTGTG 240
DB 460 GACATGAGCTGCTGCTGACTCTCAACCCGCTTCACTTCCCTCACTCCAGGCTTGTG 519
QY 241 GACCATTAC 249
DB 520 GACCATTAC 528

RESULT 4
AX443133
LOCUS AX443133 1183 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 74 from Patent WO0216599.
ACCESSION AX443133
VERSION AX443133.1 GI:21690555
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1
AUTHORS Burgess, C.E., Conley, P.B., Grose, W.M., Hart, M., Kekuda, R.,
Shinkets, R.A., Spytek, K.A., Szekeres, E.S., Tomlinson, J.E.,
Topper, J.N. and Yang, R.B.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0216599-A 74 28-FEB-2002;
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)
LOCATION/Qualifiers
1. 1183
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source
BASE COUNT 251 a 359 c 333 g 240 t
ORIGIN

Query Match 100.0%; Score 249; DB 6; Length 1183;
Best Local Similarity 100.0%; Pred. No. 1.3e-58;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGCTGATGAGGGCCCTGAGCAGGAGAAAGCAGAGAACTGCTGTTTACTCGGGAAC 60
DB 677 TGGCTGATGAGGGCCCTGAGCAGGAGAAAGCAGAGAACTGCTGTTTACTCGGGAAC 736
QY 61 CCGTAGAGGGCCCTTCTCTATCCGGGAGAGCCCAACAGAGAGGCTCTTACTCTGTCA 120
DB 737 CCGTAGAGGGCCCTTCTCTATCCGGGAGAGCCCAACAGAGAGGCTCTTACTCTGTCA 796
QY 121 GTCCGCTCAGCCGCTTCATCTCGGAGCCGATCAGACTACAGATCCACTGCTT 180

DB 797 GTCCGCTCAGCCGCTTCATCTCGGAGCCGATCAGACTACAGATCCACTGCTT 856
QY 181 GACATGAGCTGCTGCTGACTCTCAACCGGCTTCACTTCCCTCACTCCAGGCTTGTG 240
DB 857 GACATGAGCTGCTGACTCTCAACCGGCTTCACTTCCCTCACTCCAGGCTTGTG 916
QY 241 GACCATTAC 249
DB 917 GACCATTAC 925

RESULT 5
AX443135/c
LOCUS AX443135 1183 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 76 from Patent WO0216599.
ACCESSION AX443135
VERSION AX443135.1 GI:21690556
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1
AUTHORS Burgess, C.E., Conley, P.B., Grose, W.M., Hart, M., Kekuda, R.,
Shinkets, R.A., Spytek, K.A., Szekeres, E.S., Tomlinson, J.E.,
Topper, J.N. and Yang, R.B.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 0216599-A 76 28-FEB-2002;
Curagen Corporation (US); COR THERAPEUTICS, INC. (US)
LOCATION/Qualifiers
1. 1183
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 240 a 333 c 359 g 251 t
ORIGIN

Query Match 100.0%; Score 249; DB 6; Length 1183;
Best Local Similarity 100.0%; Pred. No. 1.3e-58;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGGCTGATGAGGGCCCTGAGCAGGAGAAAGCAGAGAACTGCTGTTTACTCGGGAAC 60
DB 507 TGGCTGATGAGGGCCCTGAGCAGGAGAAAGCAGAGAACTGCTGTTTACTCGGGAAC 448
QY 61 CCGTAGAGGGCCCTTCTCTATCCGGGAGAGCCCAACAGAGAGGCTTACTCTGTCA 120
DB 447 CCGTAGAGGGCCCTTCTCTATCCGGGAGAGCCCAACAGAGAGGCTTACTCTGTCA 388
QY 121 GTCCGCTCAGCCGCTTCATCTCGGAGCCGATCAGACTACAGATCCACTGCTT 180
DB 387 GTCCGCTCAGCCGCTTCATCTCGGAGCCGATCAGACTACAGATCCACTGCTT 328
QY 181 GACATGAGCTGCTGCTGACTCTCAACCCGCTTCACTTCCCTCACTCCAGGCTTGTG 240
DB 327 GACATGAGCTGCTGCTGACTCTCAACCCGCTTCACTTCCCTCACTCCAGGCTTGTG 268
QY 241 GACCATTAC 249
DB 267 GACCATTAC 259

RESULT 6
AK025645
LOCUS AK025645 2415 bp mRNA linear PRI 29-SEP-2000
DEFINITION Homo sapiens cDNA: FLJ21992 fls, clone HBP06554.
ACCESSION AK025645
VERSION AK025645.1 GI:10438227
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens hepatoma cell_line: HepG2 cDNA to mRNA, clone_1lb:HBP
clone:HBP06554.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (sites)
 AUTHORS Kawabata, A., Hiki, T., Kobatake, N., Inagaki, H., Ikema, Y.,
 Okamoto, S., Okitani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,
 Shibahara, T., Tanaka, T., Nakamura, Y., Isegai, T. and Sugano, S.
 TITLE NED0 human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2415)
 AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isegai, T.,
 Shibahara, T., Tanaka, T. and Nakamura, Y.
 TITLE Direct Submission
 JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure Analysis, Human
 Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
 Japan (E-mail: cdna1@ms.u-tokyo.ac.jp, Tel: 81-3-5449-5286,
 Fax: 81-3-5449-5416)
 COMMENT NED0 human cDNA sequencing project supported by Ministry of
 International Trade and Industry of Japan; cDNA full insert
 sequencing: Research Association for Biotechnology; cDNA library
 construction: 5'- & 3'-end one pass sequencing: Department of
 Virology and Human Genome Center, Institute of Medical Science,
 University of Tokyo (partly supported by Science and Technology
 Agency).

FEATURES
 source Location/Qualifiers
 1..2415
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HEP06554"
 /cell_line="HEP2"
 /cell_type="hepatoma"
 /clone_1b="HEP"
 /note="Cloning vector pME18SFL3"
 61..846
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB15201.1"
 /db_xref="GI:10438228"
 /translation="MGLSPKRSKSLPSLSVQGGPYTMEERKATNAVALGSP
 AGGPAELSLRGEPLTIVSEBDQMTVLSEVGREYINPSVHAKYSHMLYEGLSRE
 KAEELILPGNPGAFILRESORRGYSLSVLSRPSAMDRIHRHRIKCDNGLMYI
 SPRTFPLCALVDHYSELADICCLIKRPELOAGLPKDIPLPTVQRTPLNMK
 ELDSLLFSEATGSESLISGLRESLSYISLNDVANSDDA"

CDS
 61..846
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAB15201.1"
 /db_xref="GI:10438228"
 /translation="MGLSPKRSKSLPSLSVQGGPYTMEERKATNAVALGSP
 AGGPAELSLRGEPLTIVSEBDQMTVLSEVGREYINPSVHAKYSHMLYEGLSRE
 KAEELILPGNPGAFILRESORRGYSLSVLSRPSAMDRIHRHRIKCDNGLMYI
 SPRTFPLCALVDHYSELADICCLIKRPELOAGLPKDIPLPTVQRTPLNMK
 ELDSLLFSEATGSESLISGLRESLSYISLNDVANSDDA"

BASE COUNT 599 a 662 c 618 g 536 t
 ORIGIN

Query Match 100.0%; Score 249; DB 9; Length 2415;
 Best Local Similarity 100.0%; Pred. No. 1.2e-58;
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCGTGTATGAGGGCTGAGCAGGAGAAAGCAGGAAGTCTGTGTACTCTGGGAC 60
 DB 340 TGCGTGTATGAGGGCTGAGCAGGAGAAAGCAGGAAGTCTGTGTACTCTGGGAC 399
 QY 61 CCGTGAAGGGGCTTCTCTATCCGGGAGACGACGACGAGAGGCTTACTCTGTCTCA 120
 DB 400 CCGTGAAGGGGCTTCTCTATCCGGGAGACGACGACGAGAGGCTTACTCTGTCTCA 459
 QY 121 GTCCGCTCAGCGCCCTGATCTCGGAGACGACGACGAGAGGCTTACTCTGTCTCA 180
 DB 460 GTCCGCTCAGCGCCCTGATCTCGGAGACGACGACGAGAGGCTTACTCTGTCTCA 519
 QY 181 GACCAATGCTGTGATCATTCACCGCGCTCACTTCCCTCACTCCAGGCCCTGTG 240
 DB 520 GACCAATGCTGTGATCATTCACCGCGCTCACTTCCCTCACTCCAGGCCCTGTG 579
 QY 241 GACCAATTAC 249
 DB 580 GACCAATTAC 588

RESULT 7
 AX452880
 LOCUS
 DEFINITION Sequence 1 from Patent WO0242457.

2567 bp DNA linear PAT 06-JUL-2002

ACCESSION AX452880
 VERSION AX452880.1 GI:21712520
 KEYWORDS
 SOURCE human
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Chang, H., Yang, W. P., Wu, Y., Whitney, G. S., Perez-Villar, J. J. and
 Kanter, S. B.
 TITLE Cloning and expression of human slap-2: a novel sn2/sh3
 domain-containing human slap homologue having immune cell-specific
 expression
 JOURNAL Patent: WO 0242457-A 1 30-MAY-2002;
 Bristol-Myers Squibb Co. (US)

FEATURES
 source Location/Qualifiers
 1..2567
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 611 a 741 c 666 g 549 t
 ORIGIN

Query Match 100.0%; Score 249; DB 6; Length 2567;
 Best Local Similarity 100.0%; Pred. No. 1.2e-58;
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCGTGTATGAGGGCTGAGCAGGAGAAAGCAGGAAGTCTGTGTACTCTGGGAC 60
 DB 694 TGCGTGTATGAGGGCTGAGCAGGAGAAAGCAGGAAGTCTGTGTACTCTGGGAC 753
 QY 61 CCGTGAAGGGGCTTCTCTATCCGGGAGACGACGACGAGAGGCTTACTCTGTCTCA 120
 DB 754 CCGTGAAGGGGCTTCTCTATCCGGGAGACGACGACGAGAGGCTTACTCTGTCTCA 813
 QY 121 GTCCGCTCAGCGCCCTGATCTCGGAGACGAGATCAGACATCAGAGATCAGTGCCTT 180
 DB 814 GTCCGCTCAGCGCCCTGATCTCGGAGACGAGATCAGACATCAGAGATCAGTGCCTT 873
 QY 181 GACCAATGCTGTGATCATTCACCGCGCTCACTTCCCTCACTCCAGGCCCTGTG 240
 DB 874 GACCAATGCTGTGATCATTCACCGCGCTCACTTCCCTCACTCCAGGCCCTGTG 933
 QY 241 GACCAATTAC 249
 DB 934 GACCAATTAC 942

RESULT 8
 AF287467
 LOCUS AF287467 1348 bp mRNA linear ROD 03-JUN-2002
 DEFINITION Mus musculus Src-1-like adaptor protein-2 mRNA, complete cds.
 ACCESSION AF287467
 VERSION AF287467.1 GI:17351918
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1
 AUTHORS Loreto, M. P., Berry, D. M. and McGlade, C. J.
 TITLE Functional cooperation between c-Cbl and Src-1-like adaptor protein 2
 in the negative regulation of T-cell receptor signaling
 JOURNAL Mol. Cell. Biol. 22 (12), 4241-4255 (2002)
 MEDLINE 22022020
 PUBMED 12024036

REFERENCE 2 (bases 1 to 1348)
 AUTHORS Loreto, M. P. and McGlade, C. J.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUN-2000) Brain Tumour Research Centre, Hospital for
 Sick Children, 555 University Avenue, Toronto, Ont M5G 1X8, Canada

FEATURES
 source Location/Qualifiers
 1..1348
 /organism="Mus musculus"

CDS
/db_xref="taxon:10090"
/dev_stage="day 15 embryo"
282..1061
/note="SLAP-2"
/codon_start=1
/product="Src-1-like adaptor protein-2"
/protein_id="AA138196.1"
/db_xref="GI:17351919"

misc_feature
/note="Region: SH3 domain"
552..824
misc_feature
/note="Region: SH2 domain"
1301..1306
polyA_signal
324 a 385 c 362 g 277 t
BASE COUNT
ORIGIN

Query Match
Best Local Similarity 91.2%; Pred. No. 7.4e-49;
Matches 227; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 TGGCTGTATGAGGCGCTTGAGCGAGGAGAAAGCAGAGAACTGCTGTTTACTTGGGAAC 60
DB 558 TGGCTGTATGAGGCGCTTGAGCGAGGAGAAAGCAGAGAACTGCTGTTTACTTGGGAAC 617
QY 61 CCGGAGGCGCTTCTCTCATCCGAGAGAGCCAGACAGAGAGGCTTTACTCTGTGCA 120
DB 618 CCGGAGGCGCTTCTCTCATCCGAGAGAGCCAGACAGAGAGGCTTTACTCTGTGCA 677
QY 121 GTCCGCTCAGCGCGCTTGATCTCGGAGACCGGATCAGACATAGAGATCCACTGCTT 180
DB 678 GTCCGCTCAGCGCGCTTGATCTCGGAGACCGGATCAGACATAGAGATCCACTGCTT 737
QY 181 GACATGCGTGGCTGTATCTACCGGCGCTCACTTCCCTCACTCCAGCGCTGCTG 240
DB 738 GACATGCGTGGCTGTATCTACCGGCGCTCACTTCCCTCACTCCAGCGCTGCTG 797
QY 241 GACCTTAC 249
DB 798 GAGCATTTAC 806

RESULT 9
AF434990 1321 bp mRNA linear ROD 20-MAY-2002
LOCUS
DEFINITION Mus musculus Src-1-like adaptor protein-2 mRNA, complete cds.
ACCESSION AF434990
VERSION AF434990.1 GI:19224130
KEYWORDS
SOURCE
ORGANISM

Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
pandey,A., Ibarrola,N., Krachmarova,I., Fernandez,M.M.,
Constantinescu,S.N., Ohara,O., SawasdiKosol,S., Lodish,H.F. and
Mann,M.
TITLE
A novel Src homology 2 domain-containing molecule, Src-1-like adaptor
protein-2 (SLAP-2), which negatively regulates T cell receptor
signaling
J. Biol. Chem. 277 (21), 19131-19138 (2002)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Research, Nine Cambridge Center, Cambridge, MA 02142, USA
Location/Qualifiers

SOURCE
1..1321
/organism="Mus musculus"
/db_xref="taxon:10090"
274..1053
/note="SLAP-2; contains Src-3 and Src-2 domains"
/codon_start=1
/product="Src-1-like adaptor protein-2"
/protein_id="AA186403.1"
/db_xref="GI:19224131"

BASE COUNT 298 a 381 c 363 g 279 t
ORIGIN

Query Match
Best Local Similarity 90.8%; Pred. No. 2.1e-48;
Matches 226; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 1 TGGCTGTATGAGGCGCTTGAGCGAGGAGAAAGCAGAGAACTGCTGTTTACTTGGGAAC 60
DB 550 TGGCTGTATGAGGCGCTTGAGCGAGGAGAAAGCAGAGAACTGCTGTTTACTTGGGAAC 609
QY 61 CCGGAGGCGCTTCTCTCATCCGAGAGAGCCAGACAGAGAGGCTTTACTCTGTGCA 120
DB 610 CCGGAGGCGCTTCTCTCATCCGAGAGAGCCAGACAGAGAGGCTTTACTCTGTGCA 669
QY 121 GTCCGCTCAGCGCGCTTGATCTCGGAGACCGGATCAGACATAGAGATCCACTGCTT 180
DB 670 GTCCGCTCAGCGCGCTTGATCTCGGAGACCGGATCAGACATAGAGATCCACTGCTT 729
QY 181 GACATGCGTGGCTGTATCTACCGGCGCTCACTTCCCTCACTCCAGCGCTGCTG 240
DB 730 GACATGCGTGGCTGTATCTACCGGCGCTCACTTCCCTCACTCCAGCGCTGCTG 789
QY 241 GACCTTAC 249
DB 790 GAGCATTTAC 798

RESULT 10
HSDJ977B1/c 145068 bp DNA linear PRI 20-JUL-2001
LOCUS
DEFINITION Human DNA sequence from clone RPS-977B1 on chromosome 20. Contains
ESTs, STSs, GSSs and three putative CpG islands. Contains the 3'
end of the gene for a novel protein tyrosine kinase, a
heterogeneous nuclear ribonucleoprotein A3 pseudogene, the gene for
three isoforms of a novel protein similar to putative
RAB5-interacting protein, the TGF2 gene for TGF(beta)-induced
transcription factor 2 with two isoforms, the MYR2 gene for myosin
regulatory light chain 2 (smooth muscle isoform), the 3' end of the
gene KIA0964 (ortholog of rat PSD-95/SAP90-associated protein 4)
with two isoforms and a novel gene, complete sequence.
AI050318
AI050318.13 GI:9581785

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Lloyd,D.
TITLE
JOURNAL
COMMENT
Submitted (19-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Jul 28, 2000 this sequence version replaced gi:5924017.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (19-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Jul 28, 2000 this sequence version replaced gi:5924017.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations


```

repeat_region      /note="MIR repeat: matches 11. .236 of consensus"
11767. .11782      /note="L2 repeat: matches 2692. .2708 of consensus"
repeat_region      /note="L2083
11783. .12083      /note="AluX repeat: matches 1. .300 of consensus"
repeat_region      /note="L2121
12084. .12121      /note="L2 repeat: matches 2661. .2692 of consensus"
repeat_region      /note="L2714
12631. .12714      /note="HAI1 repeat: matches 1589. .1665 of consensus"
repeat_region      /note="L1367
13077. .13667      /note="AluUo repeat: matches 3. .290 of consensus"
repeat_region      /note="L1342
13440. .13742      /note="AluX repeat: matches 1. .304 of consensus"
repeat_region      /note="L1497
13932. .14197      /note="L1MB7 repeat: matches 5915. .6173 of consensus"
repeat_region      /note="L14512
14198. .14512      /note="AluY repeat: matches 1. .307 of consensus"
repeat_region      /note="L14756
14513. .14756      /note="L1MB7 repeat: matches 5650. .5915 of consensus"
repeat_region      /note="L15705
14789. .15705      /note="L1MB4 repeat: matches 5162. .6185 of consensus"
repeat_region      /note="L15705
15706. .16004      /note="AluX repeat: matches 1. .298 of consensus"
repeat_region      /note="L16004
16005. .16060      /note="L1MB4 repeat: matches 5111. .5162 of consensus"

Query Match      60.2%; Score 150; DB 9; Length 145068;
Best Local Similarity 86.8%; Pred. No. 2.7e-31;
Matches 165; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 60 CCTGAGAGGGGCTTCCTGATCCGGAGAGCGAGACGAGAGAGGCTTACTCTGTGC 119
Db 126136 CCCAGCAGGGCTCTGACGACCCCTGAGCCCTTCTCTCAGGCTTACTCTGTGC 126077

QY 120 AGTCGGCTTGAAGCCGCTGATCTCTGAGCCGATCAGACACTACAGATCAGCTGCT 179
Db 126076 AGTCGGCTTGAAGCCGCTGATCTCTGAGCCGATCAGACACTACAGATCAGCTGCT 126017

QY 180 TGACATAGGCTGCTGATCATCTACGCGGCTCACCCTCCCTCAGCTCAGGCGCTGT 239
Db 126016 TGACATAGGCTGCTGATCATCTACGCGGCTCACCCTCCCTCAGCTCAGGCGCTGT 125957

QY 240 GGACCATTAC 249
Db 125956 GGACCATTAC 125947

RESULT 11
AC026539/c      145833 bp      DNA      linear      HTG 27-APR-2000
LOCUS      Homo sapiens chromosome 20 clone RP11-712N14 map 20, WORKING DRAFT
DEFINITION      AC026539
SEQUENCE      AC026539.2 GI:7656813
VERSION      HTG: HTGS PHASE1; HTGS_DRAFT.
KEYWORDS      Homo sapiens.
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 145833)
AUTHORS      Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
      Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bede, P.,
      Boguslavsky, L., Bouckhagalter, B., Brown, A., Burkett, G.,
      Campiano, A., Castle, A., Choquet, Y., Colangelo, M., Collins, S.,
      Collymore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,
      Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
      Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
      Grand-Pierre, N., Grant, G., Hago, B., Heath, M., Horton, L.,
      Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatae, A.,

```

TITLE

JOURNAL

COMMENT

```

Klein, J., Labouque, K., Lamazare, R., Landers, T., Lehotzky, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGuire, A., McKernan, K., McPherson, R.,
Meldrum, J., Meneses, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,
Murphy, T., Naylor, T., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Tajamas, J.,
Testaye, S., Theodore, J., Tittel, A., Travers, M., Trisilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (22-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 27, 2000 this sequence version replaced gi:7283243.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7115
Center clone name: 712_N_14
----- Summary Statistics
Sequencing vector: M13; NT7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 125577 bases at least Q40
Consensus quality: 135703 bases at least Q30
Consensus quality: 139593 bases at least Q20
Insert size: 182000; agarose-1p
Insert size: 142233; sum-of-contigs
Quality coverage: 2.6 in Q20 bases; agarose-1p
Quality coverage: 3.3 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 37 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1214: contig of 1214 bp in length
1215 1314: gap of 100 bp
1315 2675: contig of 1359 bp in length
2674 2773: gap of 100 bp
2774 4520: contig of 1747 bp in length
4521 4620: gap of 100 bp
4621 5961: contig of 1341 bp in length
5962 6061: gap of 100 bp
6062 7719: contig of 1658 bp in length
7720 7819: gap of 100 bp
7820 9799: contig of 1980 bp in length
9800 9899: gap of 100 bp
9900 11434: contig of 1535 bp in length
11435 11534: gap of 100 bp
11535 14382: contig of 2848 bp in length
14383 14482: gap of 100 bp
14483 17148: contig of 2666 bp in length
17149 17248: gap of 100 bp
17249 18238: contig of 1891 bp in length
18240 19339: gap of 100 bp
19340 21102: contig of 1763 bp in length
21103 21202: gap of 100 bp
21203 23371: contig of 2169 bp in length
23372 23471: gap of 100 bp
23472 25782: contig of 2311 bp in length
25783 25882: gap of 100 bp

```


*	25883	28824	contig of 2942 bp in length
*	28825	28924	gap of 100 bp
*	28925	31619	contig of 2695 bp in length
*	31620	31719	gap of 100 bp
*	31720	34679	contig of 2960 bp in length
*	34680	34779	gap of 100 bp
*	34780	37527	contig of 2748 bp in length
*	37528	37627	gap of 100 bp
*	37628	40146	contig of 2519 bp in length
*	40147	40246	gap of 100 bp
*	40247	43743	contig of 3497 bp in length
*	43744	43843	gap of 100 bp
*	43844	47133	contig of 3296 bp in length
*	47134	47233	gap of 100 bp
*	47234	51023	contig of 3790 bp in length
*	51024	51123	gap of 100 bp
*	51124	54935	contig of 3812 bp in length
*	54936	55035	gap of 100 bp
*	55036	59553	contig of 4518 bp in length
*	59554	59653	gap of 100 bp
*	59654	64428	contig of 4775 bp in length
*	64429	64528	gap of 100 bp
*	64529	69211	contig of 4683 bp in length
*	69212	69311	gap of 100 bp
*	69312	72201	contig of 3590 bp in length
*	72202	73001	gap of 100 bp
*	73002	76724	contig of 3723 bp in length
*	76725	76824	gap of 100 bp
*	76825	81179	contig of 4355 bp in length
*	81180	81279	gap of 100 bp
*	81280	87009	contig of 5730 bp in length
*	87100	87109	gap of 100 bp
*	87110	90855	contig of 3746 bp in length
*	90856	90955	gap of 100 bp
*	90956	96520	contig of 5565 bp in length
*	96521	96620	gap of 100 bp
*	96621	102221	contig of 5701 bp in length
*	102322	102421	gap of 100 bp
*	102422	108293	contig of 5872 bp in length
*	108294	108393	gap of 100 bp
*	108394	116689	contig of 8236 bp in length
*	116689	116789	gap of 100 bp
*	116790	125264	contig of 8475 bp in length
*	125265	125364	gap of 100 bp
*	125365	136354	contig of 10990 bp in length
*	136355	136454	gap of 100 bp
*	136455	145833	contig of 9379 bp in length

```
misc_feature .19239
17249. /note="assembly_fragment"
19340. .21102
misc_feature /note="assembly_fragment"
21203. .23371
misc_feature /note="assembly_fragment"
23472. .25782
misc_feature /note="assembly_fragment"
25883. .28824
misc_feature /note="assembly_fragment"
28925. .31619
misc_feature /note="assembly_fragment"
31720. .34679
misc_feature /note="assembly_fragment"
34780. .37527
misc_feature /note="assembly_fragment"
37628. .40146
misc_feature /note="assembly_fragment"
40247. .43743
misc_feature /note="assembly_fragment"
43844. .47133
misc_feature /note="assembly_fragment"
47234. .51023
misc_feature /note="assembly_fragment"
51124. .54935
misc_feature /note="assembly_fragment
clone end:T7
vector_side:right"
55036. .55553
/note="assembly_fragment"
```

60.24; Score 150; DB 2; Length 145833;
 Best Local Similarity 66.84; Pred. No. 2.7e-31;
 Matches 165; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Oy	60	CCCTGAGAGGGGCTTCCTCATCCGGAGAGCCACACAGAGAGGCTCTTACTCTCTCTC	119
Db	61999	CCCGACAGGGGCTCAAGTCCACCCCTCGAGCCCTTTCTCTCAGGCTCTTACTCTCTCTC	61940
Oy	120	AATCGGCTCAAGCGCGCGCTGATCTCTGGAGACGGGATTCACACATACAGATCCGACTGCT	179
Db	61939	AATCGGCTCAAGCGCGCGCTGATCTCTGGAGACGGGATTCACACATACAGATCCGACTGCT	61880
Oy	180	TGACAAATGGTGGCTGTACATCTACACAGCGGCTTACACCTTCCGTCATCCGAGCGCTGCT	239
Db	61879	TGACAAATGGTGGCTGTACATCTACACAGCGGCTTACACCTTCCGTCATCCGAGCGCTGCT	61820
Oy	240	GGACCATTTAC	249
Db	61819	GGACCATTTAC	61810

RESULT	12
ACI25701	
TOTUS	
DEFINITION	ACI25701 168498 bp DNA linear HTG 31-JUL-2002 Rattus norvegicus clone CH230-1ZL23, *** SEQUENCING IN PROGRESS ***, 56 unordered pieces.
ACCESSION	ACI25701
VERSION	ACI25701.3 GI:22004133
KEYWORDS	HTG; HTGS PHASE1.
SOURCE	Norway rat. <i>Rattus norvegicus</i>
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; <i>Rattus</i> .
REFERENCE	1 (bases 1 to 168498)
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Altschrooks,S.L., Amaralunge,H.C., Are,J.R., Ayala,M., Banks,T., Barakita,J., Benton,jr., Bimaga,K., Blankenburg,K., Bonini,D., Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P., Buhey,C., Butch,P., Burkett,C., Burrill,K.L.B., Byrd,N.C., Carson,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdry,I., Christopoulos,C.,


```

Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,
Davis M.L., Davis C., Davy-Carroll L., Dederich D.A.,
Delaney K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H.,
Douthwaite K.J., Draper H., Dugan-Rocha S., Durkin K.J.,
Einhart C., Edgar D., Edwards C.C., Elhaj C., Escoto M.,
Fall T., Ferraguto D., Flagg N., Ford J., Foster P., Franz P.,
Gabriel A., Gao J., Garcia A., Garner T., Garza N., Gill R.,
Gorrell J.H., Guevara M., Gunaratne P., Hale S., Hamilton K.,
Harris C., Harris K., Hart M., Haylak P., Hayes A., Hernandez J.,
Hernandez O., Hodgson A., Hogue M., Hollway C., Hollins L.E.,
Homes P., Howard S., Huber J., Hulik S., Hume J., Jackson L.E.,
Jacobson B., Jia Y., Johnson R., Jolivet S., Joudan S.,
Karlssohn E., Kelly S., Khan U., King L., Koryak J., Kover C.,
Kratovic J., Kuresh A., Landry N., Leal B., Lewis L.C., Lewis L.,
Li J., Li Z., Licharge O., Lieu C., Liu J., Liu W., Lonsleged H.,
Lozado R.J., Lu X., Lucier A., Lucier R., Luna R., Ma J.,
Maheshwari M., Mapua P., Martin R., Martindale A., Martinez E.,
Massey E., Mawhinney E., McLeod M.P., Meador M., Mei G., Merzner M.,
Massey G., Miner Z., Mitchell T., Mohabbat K., Morgan M., Morris S.,
Mosier M., Neal D., Newton J., Newton N., Nguyen A., Nguyen N.,
Nguyen N., Nickerson E., Nwokenkwo S., Oguh M., Okunodu G.,
Oragunye N., Oviedo R., Pace A., Payton B., Peery J., Perez L.,
Peters L., Pickens R., Primus E., Pu L.L., Quiles M., Ren Y.,
Rivers M., Rojas A., Rojucokan I., Rolfe M., Ruiz S., Savary G.,
Scherer S., Scott G., Shan H., Shoshitani N., Sisason I.,
Sodergren E., Sonaike T., Sparks A., Stanley H., Stone H.,
Sutton A., Swalek A., Tabor P., Tamerisa A., Tamerisa K., Tang H.,
Tansey U., Taylor C., Taylor T., Telford B., Thomas N., Thomas S.,
Uemami K., Vasquez J., Vera V., Villalón D., Vinson R., Wang Q.,
Wang S., Ward-Moore S., Warren R., Washington C., Wallington S.,
Williams G., Williamson A., Wleczek R., Woodson S., Worley K.,
Wu C., Wu Y., Wu Y.F., Zhou J., Zorrilla S., Nelson D.,
Weinstock G. and Gibbs R.

Direct Submission
Unpublished
2 (bases 1 to 168498)
Worley K.C.
Direct Submission
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 168498)
Worley K.C.
Direct Submission
Submitted (31-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 30, 2002 this sequence version replaced gi:22002395.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDXR
Center clone name: CH230-12U23
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap version 0.990329
Consensus quality: 126947 bases at least Q40
Consensus quality: 134611 bases at least Q30
Consensus quality: 140704 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```

* be preserved.

```

1 1011: contig of 1011 bp in length
* 1012 1111: gap of unknown length
* 1112 2750: contig of 1639 bp in length
* 2751 2850: gap of unknown length
* 2851 4471: contig of 1621 bp in length
* 4472 4571: gap of unknown length
* 4572 5661: contig of 1090 bp in length
* 5662 5761: gap of unknown length
* 5762 7099: contig of 1337 bp in length
* 7099 7199: gap of unknown length
* 7199 8633: contig of 1434 bp in length
* 8633 9741: gap of unknown length
* 9741 9841: contig of 1009 bp in length
* 9841 11151: gap of unknown length
* 11151 11251: contig of 1310 bp in length
* 11252 13300: gap of unknown length
* 13301 13400: contig of 2049 bp in length
* 13401 15599: gap of unknown length
* 15600 16834: contig of 2199 bp in length
* 16834 16933: gap of unknown length
* 16933 19211: contig of 1134 bp in length
* 19212 19311: gap of unknown length
* 19312 21222: contig of 2278 bp in length
* 21223 21322: gap of unknown length
* 21323 22490: contig of 1168 bp in length
* 22491 22590: gap of unknown length
* 22591 23817: contig of 1127 bp in length
* 23818 25380: gap of unknown length
* 25381 25480: contig of 1463 bp in length
* 25481 26540: gap of unknown length
* 26541 28029: contig of 1060 bp in length
* 28030 28129: gap of unknown length
* 28130 29264: contig of 1389 bp in length
* 29265 29364: gap of unknown length
* 29365 31375: contig of 1135 bp in length
* 31376 31475: gap of unknown length
* 31476 33952: contig of 2011 bp in length
* 33953 34052: gap of unknown length
* 34053 35294: contig of 2477 bp in length
* 35295 35394: gap of unknown length
* 35395 37412: contig of 1242 bp in length
* 37413 39442: gap of unknown length
* 39443 39542: contig of 2017 bp in length
* 39543 40751: gap of unknown length
* 40752 40851: contig of 1931 bp in length
* 40852 42914: gap of unknown length
* 42915 44652: gap of unknown length
* 44653 44752: contig of 2064 bp in length
* 44753 46734: gap of unknown length
* 46735 46834: contig of 1638 bp in length
* 46835 48906: gap of unknown length
* 48907 51535: contig of 1982 bp in length
* 51536 51635: gap of unknown length
* 51636 54767: gap of unknown length
* 54768 57371: contig of 3031 bp in length
* 57372 59651: gap of unknown length
* 59652 63033: contig of 2505 bp in length
* 63034 63133: gap of unknown length
* 63134 65285: contig of 3152 bp in length
* 65286 65385: gap of unknown length
* 65386 66573: contig of 2188 bp in length
* 66574 68574: gap of unknown length

```

REFERENCE 1 (bases 1 to 114096)
AUTHORS, Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.

- * NOTE: Estimated insert size may differ from sequence length
- * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
- * NOTE: This is a 'working draft' sequence. It currently
- * consists of 33 contigs. The true order of the pieces

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information -----
 Center project name: L7115
 Center clone name: 712 N 14
 ----- Summary Statistics -----
 Sequencing vector: M13, M77815, 100% of reads
 Chemistry: Dye-terminator Big Dye, 100% of reads
 Assembly program: Phrap, version 0.960731
 Consensus quality: 125577 bases at least Q40
 Consensus quality: 135703 bases at least Q30
 Consensus quality: 139593 bases at least Q20
 Insert size: 182000; agarose-fp
 Insert size: 142233; sum-of-ctnigs
 Quality coverage: 2.6 in Q20 bases; agarose-fp
 Quality coverage: 3.3 in Q20 bases; sum-of-ctnigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 37 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1 1214: contig of 1214 bp in length
 * 1215 1314: gap of 100 bp
 * 1315 2673: contig of 1359 bp in length
 * 2674 2773: gap of 100 bp
 * 2774 4520: contig of 1747 bp in length
 * 4521 4620: gap of 100 bp
 * 4621 5961: contig of 1341 bp in length
 * 5962 6061: gap of 100 bp
 * 6062 7719: contig of 1658 bp in length
 * 7720 7819: gap of 100 bp
 * 7820 9799: contig of 1980 bp in length
 * 9800 9899: gap of 100 bp
 * 9900 11434: contig of 1535 bp in length
 * 11435 11534: gap of 100 bp
 * 11535 14382: contig of 2848 bp in length
 * 14383 14482: gap of 100 bp
 * 14483 17148: contig of 2666 bp in length
 * 17149 17248: gap of 100 bp
 * 17249 19239: contig of 1991 bp in length
 * 19240 19339: gap of 100 bp
 * 19340 21102: contig of 1763 bp in length
 * 21103 21202: gap of 100 bp
 * 21203 23371: contig of 2169 bp in length
 * 23372 23471: gap of 100 bp
 * 23472 25782: contig of 2311 bp in length
 * 25783 25882: gap of 100 bp
 * 25883 28824: contig of 2942 bp in length
 * 28825 28924: gap of 100 bp
 * 28925 31619: contig of 2695 bp in length
 * 31620 31719: gap of 100 bp
 * 31720 34679: contig of 2960 bp in length
 * 34680 34779: gap of 100 bp
 * 34780 37527: contig of 2748 bp in length
 * 37528 37627: gap of 100 bp
 * 37628 40146: contig of 2519 bp in length
 * 40147 40246: gap of 100 bp
 * 40247 43743: contig of 3497 bp in length
 * 43744 43843: gap of 100 bp
 * 43844 47133: contig of 3290 bp in length
 * 47134 47233: gap of 100 bp
 * 47234 51023: contig of 3790 bp in length
 * 51024 51123: gap of 100 bp
 * 51124 54935: contig of 3812 bp in length
 * 54936 55035: gap of 100 bp
 * 55036 59553: contig of 4518 bp in length
 * 59554 59653: gap of 100 bp

FEATURES
 source
 * 59654 64428: contig of 4775 bp in length
 * 64429 64528: gap of 100 bp
 * 64529 69211: contig of 4683 bp in length
 * 69212 69311: gap of 100 bp
 * 69312 72901: contig of 3590 bp in length
 * 72902 73001: gap of 100 bp
 * 73002 76724: contig of 3723 bp in length
 * 76725 76824: gap of 100 bp
 * 76825 81179: contig of 4355 bp in length
 * 81180 81279: gap of 100 bp
 * 81280 87009: contig of 5730 bp in length
 * 87010 87109: gap of 100 bp
 * 87110 90855: contig of 3746 bp in length
 * 90856 90955: gap of 100 bp
 * 90956 96520: contig of 5565 bp in length
 * 96521 96620: gap of 100 bp
 * 96621 102321: contig of 5701 bp in length
 * 102322 102421: gap of 100 bp
 * 102422 108293: contig of 5872 bp in length
 * 108294 108393: gap of 100 bp
 * 108394 116689: contig of 8296 bp in length
 * 116690 116789: gap of 100 bp
 * 116790 125284: contig of 8475 bp in length
 * 125285 125364: gap of 100 bp
 * 125365 136354: contig of 10990 bp in length
 * 136355 136454: gap of 100 bp
 * 136455 145833: contig of 9379 bp in length.
 Location/Qualifiers
 1 .145833
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="20"
 /map="20"
 /clone="RP11-712N14"
 /clone_1fb="RPC1-11 Human Male 8AC"
 1 .1214
 /note="assembly_fragment"
 1315 .2673
 /note="assembly_fragment"
 2774 .4520
 /note="assembly_fragment"
 4621 .5961
 /note="assembly_fragment"
 6062 .7719
 /note="assembly_fragment"
 7820 .9799
 /note="assembly_fragment"
 9900 .11434
 /note="assembly_fragment"
 11535 .14382
 /note="assembly_fragment"
 14483 .17148
 /note="assembly_fragment"
 17249 .19239
 /note="assembly_fragment"
 19340 .21102
 /note="assembly_fragment"
 21203 .23371
 /note="assembly_fragment"
 23472 .25782
 /note="assembly_fragment"
 25883 .28824
 /note="assembly_fragment"
 28925 .31619
 /note="assembly_fragment"
 31720 .34679
 /note="assembly_fragment"
 34780 .37527
 /note="assembly_fragment"
 37628 .40146
 /note="assembly_fragment"
 40247 .43743
 /note="assembly_fragment"

misc_feature 43844..47133
/note="assembly_fragment"
misc_feature 47234..51023
/note="assembly_fragment"
misc_feature 51124..54935
/note="assembly_fragment"
clone_end:T7
vector_side:right"
55036..55953
/note="assembly_fragment"

Query Match 42.9%; Score 106.8; DB 2; Length 145833;
Best Local Similarity 94.1%; Pred. No. 2.7e-19;
Matches 111; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TGGCTGATAGAGGCGCTTCCATCCGGAGAGCCAGAGAGGCTCTTACTCTCTGT 118
|||
Db 132536 TGGCTGATAGAGGCGCTTCCATCCGGAGAGCCAGAGAGGCTCTTACTCTCTGT 118
|||
Db 132536 CTTGAGAGGCGCTTCCATCCGGAGAGCCAGAGAGGCTCTTACTCTCTGT 132653

RESULT 15

HUMHCKB 1926 bp mRNA linear PRI 08-NOV-1994
LOCUS Human hemopoietic cell protein-tyrosine kinase (HCK) gene, complete
DEFINITION cds, clone HK24.
ACCESSION M16592
VERSION M16592.1 GI:183913
KEYWORDS kinase; protein kinase; protein-tyrosine kinase.
SOURCE Human mitogen-stimulated leukocyte, CDNA to mRNA, clone HK24.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1926)
AUTHORS Ziegler, S.F., Marth, J.D., Lewis, D.B. and Perlmuter, R.M.
TITLE Novel protein-tyrosine kinase gene (hck) preferentially expressed
in cells of hemopoietic origin
JOURNAL Mol. Cell. Biol. 7 (6), 2276-2285 (1987)
MEDLINE 87257943
PUBMED 3453117

FEATURES

source Location/Qualifiers
1..1926
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="20q11-q12"
1..1926
/gene="HCK"
/gene="HCK"
/gene="HCK"
/product="HCK mRNA"
76..1593
/gene="HCK"
/note="protein-tyrosine kinase"
/codon_start=1
/protein_id="AA52644.1"
/db_xref="GI:306833"
/translation="MGCMKSKFLQYGVNFTSKTETSASPNCVYVPPDPTSTIKPNS
HNSNTPGIRAGSEDIIVVALYDAIHHEDLSFOKGDQVVLSESGEMKASLATR
KEGYIPSNVAVRDLSETEEMFFGSRKDAEROLAPGMNLSFMRDSEETKGSYS
LSVRDPRGDDTVKHKYKIRTLNDNGFYISPRSTLQELVDHKKNGDLQQLTSLV
PMSKPKQPKREKDAWEIPRESLKLEKLAGQFGEVMATYKNTKYAVKTKPESM
SVEAFABANVMKTLQHDKLVLAHVATKEPITITTEPMAGSLDLFLKSDSGKQPL
PFLIDFSQIAEGNAFIEQRNYIHRDLRAANILVSAIVCKIADFLARVIDNEVTA
REGAFPIKWTABEAINFGSFTIKSDVSGFILLMEIVYGRIPYPMGNSPEVITALE
RGYMRPENCPEBELYNIMRCMKRPERPTEFYIOSVLDPYATATESQYCOOP"

BASE COUNT 497 a 522 c 520 g 387 t
ORIGIN 1 bp upstream of EcoRI site; chromosome 20q11-q12.

Query Match

40.2%; Score 100.2; DB 9; Length 1926;

Best Local Similarity 62.7%; Pred. No. 2.6e-17;
Matches 156; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

Qy 1 TGGCTGATAGAGGCGCTTCCATCCGGAGAGCCAGAGAGGCTCTTACTCTCTGT 60
|||
Db 442 TGGTTTTTCAAGGCGATAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 501
|||
Qy 61 CTTGAGAGGCGCTTCCATCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
|||
Db 502 ATGCTGGGCTCTTCCATGATCCGGATVAGCGAGACACTAAGAGAGCTACTTGTCC 561
|||
Qy 121 GTCCGCTCAGCGCGCTTCCATCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
|||
Db 562 GTGCGAGACTAGACCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 621
|||
Qy 181 GACAAAGCTGGCTGTATCTCAACCGGCGCTCACTTCCCTCACTCCAGGCGCTGTG 240
|||
Db 622 GACACGCGGCGCTTCTATCATATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 681
|||
Qy 241 GACCATTTAC 249
|||
Db 682 GACCACTAC 690

Search completed: March 30, 2003, 06:27:55
Job time : 903.77 secs

